



Repetition and Reputation: Implications for Trust and Trustworthiness When Institutions Change

Citation

Bohnet, Iris, and Steffen Huck. 2004. "Repetition and Reputation: Implications for Trust and Trustworthiness When Institutions Change." *American Economic Review* 94 (2): 362–66. <https://doi.org/10.1257/0002828041301506>.

Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:37894215>

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

Share Your Story

The Harvard community has made this article openly available. Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

REALISM IN EXPERIMENTAL ECONOMICS: INTEGRATING THE REAL WORLD INTO EXPERIMENTS

Repetition and Reputation: Implications for Trust and Trustworthiness When Institutions Change

By IRIS BOHNET AND STEFFEN HUCK*

Institutions change—in Iraq, the formerly Communist countries, and many private and public organizations. But do people adapt to the new institutional environment, and if so, how quickly? This paper examines institutional change—how long it takes people to transition from one institutional environment to another or, put differently, whether old institutional regimes have an afterglow. More specifically, we study whether trust and trustworthiness can be fostered by first exposing people to an environment conducive to trust. We are interested in whether (intrinsic) trust and trustworthiness can be induced in the long run by providing extrinsic incentives for trust and trustworthiness in the short run.

Reputation systems may provide incentives for trustworthiness and trust. Direct reputation-building may occur in repeated games where pairs of subjects play the same stage game repeatedly, but repeat transactions are not necessarily the rule in today's global economy. In population games where agents are randomly re-matched in every period, indirect reputation systems are a potential substitute for personal interactions—provided information about others' past behavior is available. On eBay, for example, buyers are willing to pay a premium

of 8.1 percent of the selling price to a seller with an established good reputation (Paul Resnick et al., 2003).

This paper examines experimentally to what degree indirect reputation-building substitutes for direct reputation-building in repeat interactions in the short run and analyzes the effects these environments have on behavior in the long run. In contrast, most earlier experimental studies focus on one-shot and repeat interactions in the short run.¹

We compare the effects of direct and indirect reputation-building in a binary-choice trust game where a buyer (the trustor) can either interact with the seller (the trustee) or exit. The trustee can either honor or exploit trust. The payoffs are such that a money-maximizing trustee prefers exploiting to honoring trust in a one-shot game, while a money-maximizing trustor prefers not offering trust to being exploited. The unique Nash equilibrium of the single-shot game predicts no trade. Figure 1 presents the game we implemented with the actual payoffs in cents used.

In our experiment, subjects participate in the trust game in two blocks of 10 rounds each, which is common knowledge.² In phase 1, the first 10 rounds, they are confronted either with a standard, “one-shot” random matching treatment (“stranger” or “S”); a fixed-pairs, finitely repeated game treatment (“partner” or “P”); or

* Bohnet: Kennedy School of Government, Harvard University, 79 JFK Street, Cambridge, MA 02139 (e-mail: Iris_Bohnet@Harvard.edu); Huck: Department of Economics and ELSE, University College London, Gower Street, London WC1E 6BT, United Kingdom (e-mail: s.huck@ucl.ac.uk). We thank Rachel Croson and the participants at the 2004 ASSA meetings for their helpful comments and Jeffrey Bielicki for his excellent research assistance. We gratefully acknowledge financial support from the Russell Sage Foundation, the Leverhulme Trust, and the Economic and Social Research Council (U.K.) via ELSE.

¹ For a recent survey, see James Andreoni and Rachel Croson (2004). Studies examining the effects of different institutional environments over time include Bohnet et al. (2001) and Ernst Fehr and Simon Gächter (2003).

² The experiments were computerized using Urs Fischbacher's (1999) z-tree software. The instructions are available upon request.

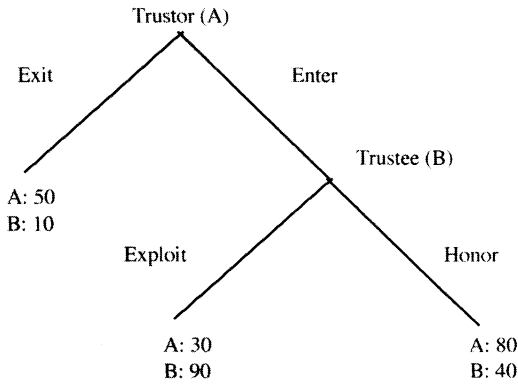


FIGURE 1. THE TRUST GAME

TABLE 1—TRUST AND TRUSTWORTHINESS RATES IN PHASE 1 (ROUNDS 1–10) AND PHASE 2 (ROUNDS 11–20)

Treatment	Trust rate		Trustworthiness rate	
	Phase 1	Phase 2	Phase 1	Phase 2
S: stranger ($N = 48$)	0.32	0.23	0.30	0.18
RS: reputation-stranger ($N = 57$)	0.43	0.19	0.55	0.18
P: partner ($N = 51$)	0.59	0.32	0.61	0.28

In this paper, we take an empirical approach and examine whether there are any *history effects* and, if so, whether they are systematic. In particular, we estimate subjects' propensity to trust (or to be trustworthy) in the second phase of the experiment as a function of the institution they were exposed to in the first phase, their experiences in the first and second phase, their type (as measured by their initial propensity to trust and be trustworthy), and time. In our data, we find that subjects do understand changes in the incentive structure and fully discount previous experiences if they were not gained in the same environment. While this is in line with orthodox theory, our second main finding challenges it. We find that exposure to a partner treatment makes trustees more trustworthy in the long run. Partner and reputation-stranger treatments produce similar results in the short but not in the long run.

I. Experimental Results

Table 1 presents average trust and trustworthiness rates for each treatment and the two phases of the experiment (Figs. 2 and 3 present the data by round). Trust rates indicate the fraction of trustors offering trust in a given round; trustworthiness rates indicate the fraction of trustees honoring trust in a given round, conditional on having been offered trust.

In phase 1, trustors are significantly more likely to trust in P than in S, and somewhat more likely to trust in RS than in S. Trustees are equally likely to honor trust in P and RS and significantly less likely to do so in S.³ Our results show the existence of

a random-matching treatment (“reputation-stranger” or “RS”). In the latter, trustors are informed about their trustees' past behavior in each round. In phase 2, rounds 11–20, all subjects interact in the stranger environment (without information about the past). In all, 312 subjects participated in our experiment; 96 in the S-treatment (four sessions), 102 in the P-treatment (five sessions), and 114 in the RS-treatment (four sessions). Roles were randomly assigned and kept fixed during the experiment.

For the short run (phase 1), models incorporating incomplete information about agents' preferences and/or rationality allow for reputation-building, directly and indirectly, in finitely repeated games (David Kreps et al., 1982). If there are such *reputation effects*, there should be more trust and trustworthiness in P and RS than in S in the short run and a decrease of trust and trustworthiness toward the end of the first 10 rounds. Our results for the first phase are in line with this prediction.

For the long run (phase 2), all orthodox models predict the same behavior. We should see low (or zero) levels of trust and trustworthiness in all three treatments since incentives for building a reputation have been removed. In that sense, orthodox models predict that history does not matter. If, on the other hand, reputation-based interactions “crowd in” trust and trustworthiness or evoke specific norms of behavior, differences between the treatments might be observed. Theoretically, such long-term effects require either changes in preferences (Bohnet et al., 2001) or some inertia in adjustment and learning (Ido Erev and Alvin Roth, 1998).

³ Mann-Whitney U tests using session averages aggregated over the first 10 rounds as independent observations

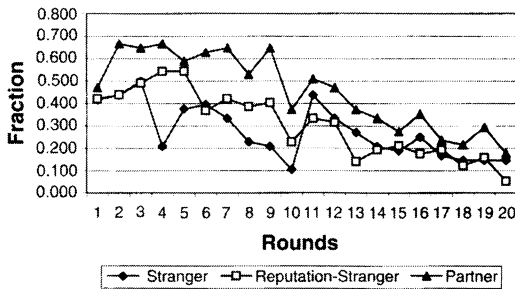


FIGURE 2. TRUST RATES

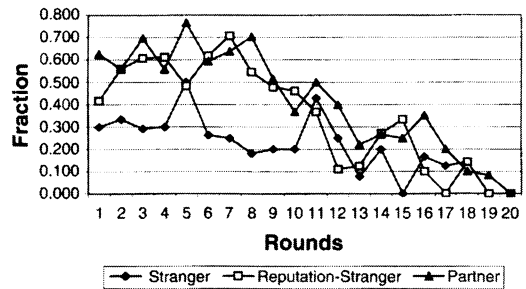


FIGURE 3. TRUSTWORTHINESS RATES

rather strong reputation effects, especially for trustees.

In phase 2, trust and trustworthiness rates in P are slightly higher than in the two stranger treatments, suggesting history effects. The differences between P and RS are significant.⁴ To get a clearer picture of subjects' behavior in the second phase we estimate linear probability models for trustors' propensity to trust and trustees' propensity to be trustworthy.⁵ We run two random-effects panel regressions:

$$\begin{aligned}
 \text{RATE}_{i,t} = & \alpha + \beta \text{RS}_i + \gamma P_i + \delta \text{EFP}_i \\
 & + \varepsilon \text{EFP}_i \times \text{RS} + \zeta \text{EFP}_i \times P \\
 & + \eta \text{ESP}_{i,t} + \theta \text{ESP}_{i,t} \times \text{RS} \\
 & + \iota \text{ESP}_{i,t} \times P + \kappa \text{TYPE}_i \\
 & + \lambda \text{TYPE}_i \times \text{RS} + \mu \text{TYPE}_i \times P \\
 & + \nu \text{RD}_i + \xi \text{RD}_i \times \text{RS} + \omicron \text{RD}_i \times P \\
 & + v_i + e_{i,t}
 \end{aligned}$$

reveal the following p values (two-tailed) for trust: P-S, $p = 0.01$; RS-S, $p = 0.08$; P-RS, $p = 0.14$. For trustworthiness, the corresponding p values were: P-S, $p = 0.01$; RS-S, $p = 0.02$; P-RS, $p = 0.33$.

⁴ Mann-Whitney U tests using session averages aggregated over the second 10 rounds as independent observations reveal the following p values (two-tailed) for trust: P-S, $p = 0.22$; RS-S, $p = 0.56$; P-RS, $p = 0.03$. For trustworthiness, the corresponding p values were: P-S, $p = 0.12$; RS-S, $p = 0.56$; P-RS, $p = 0.09$.

⁵ For samples of this size, linear probability models are more robust than logit or probit models.

where $\text{RATE}_{i,t}$ is subject i 's probability to trust (in the first regression) or be trustworthy (in the second regression); α is the constant; RS_i is a dummy variable equal to 1 if subject i is in the reputation-stranger treatment and zero otherwise; P_i is a dummy variable equal to 1 if subject i is in the partner treatment and zero otherwise; EFP_i is subject i 's experience during the first phase, rounds 1–10 (i.e., a trustor's experienced trustworthiness rate or a trustee's experienced trust rate); $\text{ESP}_{i,t}$ is subject i 's experience up to period $t - 1$ in the second phase, rounds 11–20 (experienced trustworthiness rate for the trustor and experienced trust rate for the trustee up to $t - 1$); TYPE_i captures subject i 's initial propensity to trust or be trustworthy (a dummy variable for the trustor equal to 1 if the subject trusted in round 1 and 0 otherwise; and the average trustworthiness rate of subject i in the first 10 rounds for the trustee);⁶ RD_i is the round; v_i the idiosyncratic random-effect of subject i , and $e_{i,t}$ the error term. Table 2 shows the results.

For trustors we find no main treatment effects. The experience of trustworthiness in the first phase (EFP) and a subject's initial propensity to trust (TYPE) only matter in S (without institutional change). In the other treatments where trustworthiness can be strategic in the first phase, neither others' nor own actions in the first phase affect trust in the nonstrategic second phase. In contrast, recent experiences in the previous rounds of phase 2 (ESP) are important in all treatments. The more trustworthiness subjects have recently experienced, the

⁶ Taking a similar average for trustors does not make sense, since trust in later rounds may be driven by experience.

TABLE 2—ESTIMATED TRUST AND TRUSTWORTHINESS RATES IN PHASE 2 (ROUNDS 11–20)

Variable	Trust	Trustworthiness
Reputation-stranger	−0.098 (0.161)	0.185 (0.389)
Partner	0.212 (0.168)	0.652 [†] (0.364)
EFP (experience in first phase)	0.309 [†] (0.172)	−0.355 (0.317)
EFP × reputation-stranger	−0.226 (0.212)	0.019 (0.419)
EFP × partner	−0.226 (0.211)	−0.151 (0.360)
ESP (experience in second phase up to $t - 1$)	0.444** (0.109)	−0.483 (0.682)
ESP × reputation-stranger	0.020 (0.144)	0.418 (0.935)
ESP × partner	−0.208 (0.130)	0.984 (0.801)
Type	0.153* (0.076)	0.492** (0.144)
Type × reputation-stranger	−0.142 (0.101)	−0.485* (0.217)
Type × partner	0.101 (0.103)	−0.522* (0.230)
Round	−0.031** (0.006)	−0.007 (0.018)
Round × reputation-stranger	−0.008 (0.008)	−0.009 (0.025)
Round × partner	−0.005 (0.008)	−0.034 (0.024)
Constant	0.564** (0.114)	0.029 (0.277)
Number of observations:	1,380	289
Number of subjects:	138	119
R^2 :	0.171	0.130

Notes: The table presents results from linear probability regressions. Standard errors are reported in parentheses.

[†] Statistically significant at the 10-percent level.

* Statistically significant at the 5-percent level.

** Statistically significant at the 1-percent level.

more willing they are to trust.⁷ There is a significant time trend in all treatments; subjects become less trusting as the end nears.

For trustees we find that subjects in P are substantially more trustworthy than subjects in RS and a little more trustworthy than those in S. The experience of trust does not matter for

⁷ Since one can only experience trustworthiness by trusting (which might cause ESP and trust rates to be correlated) we ran two control regressions, one for the initially trusting type (TYPE = 1) and one for the skeptical type (TYPE = 0). As in the overall regression (Table 2), the coefficients for ESP are around 0.4 and highly significant.

trustworthiness, independent of whether the experience was gained in the first or the second phase (EFP or ESP). Trust does not breed trustworthiness. A subject's propensity to be trustworthy in the first phase (TYPE) is only relevant when it was nonstrategic and, thus, a true matter of type. While Figure 3 suggests that trustworthiness decreases over time, the regressions show that this is an artifact of the matching—contrary to what trustors seem to expect.

II. Discussion and Conclusions

Direct and indirect reputation systems increase trust and trustworthiness in the short run. Subjects strongly respond to the direct reputation-building opportunities in a repeated game. With indirect reputation-building, trustees appear to respond more strongly to the institutional environment than trustors. The benefits of this more complex and less familiar environment may not be as obvious as the advantages of repeat interactions, which may help explain why many consumers do not trust internet-based transactions using indirect reputation systems such as on eBay (Peter Kollock, 1999). Our results suggest that they may be too pessimistic.

We also find that past experience is more relevant for trustors than for trustees. Experiences of trustworthiness increase the likelihood of trust in the same treatment; experiences of trust have no effect on the likelihood of trustworthiness; that is, trust is not self-fulfilling as suggested, for example, by Gerardo Guerra et al. (2002).

Finally, and perhaps most importantly, we find that there are significant history effects. In particular, we find that trustees are more trustworthy after having been exposed to a partner treatment. The partner treatment is the most effective institutional arrangement to foster trust and trustworthiness in the short *and* in the long run. Indeed, it appears as if experiencing the intimate partner relationship breeds genuine trustworthiness. This might have important implications for issues in institutional design and education. Interactions in small closely knit groups may have long-lasting beneficial consequences.

REFERENCES

Andreoni, James and Croson, Rachel. "Partners versus Strangers: Random Rematching in

- Public Goods Experiments,” in V. Smith and C. Plott, eds., *Handbook of experimental economic results*. 2004 (forthcoming).
- Bohnet, Iris; Frey, Bruno S. and Huck, Steffen.** “More Order with Less Law: On Contract Enforcement, Trust, and Crowding.” *American Political Science Review*, March 2001, 95(1), pp. 131–44.
- Erev, Ido and Roth, Alvin E.** “Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria.” *American Economic Review*, September 1998, 88(4), pp. 848–81.
- Fehr, Ernst and Gächter, Simon.** “Do Incentive Contracts Crowd Out Voluntary Cooperation?” Working paper, Institute for Empirical Research in Economics, University of Zurich, Switzerland, 2003.
- Fischbacher, Urs.** “zTree: A Toolbox for Ready-made Economic Experiments.” Working paper, Institute for Empirical Research in Economics, University of Zurich, Switzerland, 1999.
- Guerra, Gerardo A.; Bacharach, Michael and Zizzo, Daniel John.** “Is Trust Self-Fulfilling? An Experimental Study.” Working paper, University of Oxford, Oxford, U.K., 2002.
- Kollock, Peter.** “The Production of Trust in Online Markets,” in E. J. Lawler, M. W. Macy, S. R. Thye, and H. A. Walker, eds., *Advances in group processes*. Greenwich, CT: JAI Press, 1999, pp. 99–123.
- Kreps, David; Milgrom, Paul; Roberts, John and Wilson, Robert.** “Rational Cooperation in the Finitely Repeated Prisoners’ Dilemma.” *Journal of Economic Theory*, August 1982, 27(2), pp. 245–52.
- Resnick, Paul; Zeckhauser, Richard; Swanson, John; and Lockwood, Kate.** “The Value of Reputation on e-Bay: A Controlled Experiment.” Working paper, Kennedy School of Government, Harvard University, 2003.